January 20: Gad Getz, FireCloud: The Future of Cancer Genome Analysis

SYNOPSIS:

The cost of DNA sequencing has dropped more than one-million-fold over the last decade, making it increasingly possible to discover the genetic basis of cancer and response to treatment. Three challenges impede this goal: 1) Analysts lack the resources to download, store and compute on the data; 2) Existing tools and infrastructure have not been designed to scale to handle petabytes or exabytes; and 3) Collaboration is hindered by the current model of storing data locally.

The large-scale sequencing efforts of TCGA has begun to elucidate the genetic pathogenesis of cancer, enabling the development of targeted therapies. However, to enter an era of true “precision medicine,” we need to create sophisticated information technologies to store, analyze, and share data. FireCloud offers a solution to these needs.

FireCloud democratizes data access and facilitates collaboration by providing a robust, scalable platform accessible to the community at large. Using the elastic compute capacity of Google Cloud, FireCloud empowers analysts, tool developers and production managers to perform large-scale analysis, engage in data curation, and store or publish results. FireCloud is modeled after Firehose, an analysis on-premesis infrastructure developed by the Getz Lab at the Broad Institute.

As in Firehose, workspaces are central to the FireCloud architecture. Workspaces are computational sandboxes that enable users to organize genomic data and metadata into a data model, run analysis methods, and view results. Users can upload their own analysis methods to workspaces or import the Broad institute’s best practice tools and pipelines. FireCloud will include tutorial workspaces, and carefully curated Open and Controlled Access TCGA workspaces which users can clone.

FireCloud will enable the mission of TCGA and other cancer genome projects by provisioning workspaces with curated data and best practice tools and pipelines. This will empower researchers across the globe to explore the TCGA data in new and innovative ways which will increase opportunities to novel contributions to cancer research.

Session details...**Bio:**

Dr. Getz is an internation-ally acclaimed leader in cancer genome analysis and is pioneering widely used analytic programs in cancer genomic sequence analysis. Dr. Getz is an Associate Professor of Pathology at Harvard Medical School. He is faculty and Director of Bioinformatics at the Massachusetts General Hospital Cancer Center and Department of Pathology, and is an Institute Member of the Broad Institute of Harvard and MIT, where he directs the Cancer Genome Computational Analysis Group. He is also the inaugural incumbent of the Paul C. Zamecnik Chair in Oncology at the MGH Cancer Center.

Research & Clinical Interests

The Getz Laboratory specializes in cancer genome analysis which includes two major steps: (i) Characterization – cataloging of all genomic events and the mechanisms that created them during the clonal evolution of the cancer, comparing events at the DNA, RNA and protein levels between tumor and normal samples from an individual patient; and (ii) Interpretation – analysis of the characterization data across a cohort of patients with the aim of identifying the alterations in genes and pathways that cause cancer or increase its risk as well as identifying molecular subtypes of the disease, their markers and relationship to clinical variables.

In addition to his role at the Broad, Getz is a co-principal investigator in the Genome Data Analysis Center (GDAC) of the NCI/NHGRI TCGA (The Cancer Genome Atlas) project; a co-leader of the International Cancer Genome Consortium (ICGC) Pan-Cancer Analysis of Whole Genomes (PCAWG) project; a co-principal investigator of the Broad-led NCI Cloud Pilot; and a member of various NCI advisory committees. Dr. Getz has published numerous papers in recent years in prominent journals that describe new methodologies to study cancer genomes, identifying new genes and pathways involved in different tumor types, mutational signatures, and tumor evolution.

Getz received his B.S. degree in physics and mathematics from Hebrew University and a M.Sc. in physics from Tel-Aviv University. He later earned a Ph. D. in physics from the Weizmann Institute of Science in Israel. He completed his postdoctoral training at the Broad Institute of Harvard and MIT with Todd Golub, where he focused on developing computational tools and analyzing expression of miRNAs across cancer.

SUMMARY:

Topic: FireCloud: The Future of Cancer Genome Analysis

Speaker: Gad Getz, Ph.D.

Date: Wednesday, January 20, 2016

Time: 11 AM – 12 PM ET

You are invited to listen to Dr. Getz's presentation in Room 2W910-912 in the NCI Shady Grove Building on Medical Center Drive or via WebEx.

Presentation: A screencast of the presentation will be available for viewing after the event on the NCI CBIIT Speaker Series YouTube Playlist

View the presentation slides.

About the NCI CBIIT Speaker Series:

The National Cancer Institute (NCI) Center for Biomedical Informatics and Information Technology (CBIIT) Speaker Series is a bi-weekly knowledge-sharing forum featuring both internal and external speakers on topics of interest to the biomedical informatics and research communities. For additional information, including past speaker series presentations, visit the CBIIT Speaker Series page.

Individuals with disabilities who need reasonable accommodation to participate in this program should contact the Office of Space and Facilities Management (OSFM) at 240-276-5900 or the Federal TTY Relay number 1-800-877-8339.